Date: Fri, 22 Aug 1997 08:25:32 -1000
From: Art Pollard <pollarda@hawaii.edu>
X-Sender: pollarda@uhunix4
To: Ted Nelson <ted@glasswings.com.au>
Subject: Re: *"Embedded Markup Considered Harmful (preview)
MIME-Version: 1.0

Ted:

The reason for my asking the question is that I have built a system which seeks to fulfill Xanadu's design goals. (Since the Web and every other hypertext system is lame IMHO.)

Currently, the system will take the text and compile it into a file which contains the text, the index for the text, and any meta information such as graphics, etc. This data file is approximately 60% of the size of the original ASCII text. And this also includes a unique ID for every byte.

In addition, the whole thing is dynamic so new text may be added and old text may be deleted. (This part isn't finished yet.)

The only embeded tags which the system requires are tags to break up the text into records and other controls for the indexing engine. So, for example, you would want to put a record break tag between paragraphs.

The system is designed to support autonomous links which are completely searchable and have attributes such as author, creation date, subject, etc.

So, that is why I am interested in some of the issues surrounding transclusion.

To be frank, I hadn't thought of textual attributes being treated as links as well. Though it makes perfect sense.

Of couse, the Web can't to anything approaching the goals of Xanadu. One of the things which the Web does push is the structural markup of text. I can see how this can be quite powerful though not necissarily for the reasons that the web people have stated. The SGML crowd seems to always push the idea of searching for text within a certain textual structure. While I can see why someone might want to search for a certain set of words in a header, or paragraph, etc., in practice nobody does so and text retrieval techniques (such as ranking) are such that it probably wouldn't make much difference anyway to the search results to further narrow the

→ SYSTEM WITH 'AUTONOMOUS LINKS'

search by specifing header, paragraph, etc.. (Generally what you are looking for is listed in the first couple of hits which is why it doesn't make much difference.) Furthermore, there seems to be no correspondence between SGML tags as they are used and names that non-computer-geeks can remember. So for example, if someone _did_ want to find something in a header, how would they know that a HTML header is H1 through H6? If there were a means of identifying a "Header Level 1" with "H1" as far as naming is concerened, SGML based languages would be much more useful for computer novices. IMHO, SGML type markup is nice for textual processsing, and automated formatting but really doesn't lend itself to web publishing well.

One nice thing which has come of late is the idea of using style sheets to specify the markup for a particular textual structure. This does make a lot of sense in many ways. However, it still falls short of Xanadu's goals.

Well, this brings me to the point of this letter. What do you think about the idea of attatching styles to particular textual structures as defaults and then supporting custom styles/style overrides as links? The advantage of this is that there would be a significantly reduced space overhead for styles as well as a greatly increased speed of delivery as not so many styles would have to be "served" out of the style/link database.

Thank you for your time and feedback,

-Art

[BTW: I am greatly looking forward to getting this project done and people using it. However, right now, there is only a couple of us working on it and we are looking for funding so we can devote more time as well as bring on additional people. Cheers!]

Art Pollard <PollardA@Hawaii.edu>
Moderator for Comp.Theory.Info-Retrieval
List Maintainer for the Hyper-Theory (Hypertext Theory) mailing list.